



Mobilité urbaine: Apprentissage automatique pour la construction de simulateurs à l'aide de masses de données

**Co-Directeurs**

*Jean-Marc Jézéquel (DiverSE)*

*David Gross-Amblard (DRUID)*

**Responsable scientifique**

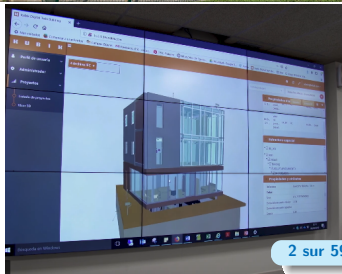
*Sébastien Leparoux*

**Doctorant**

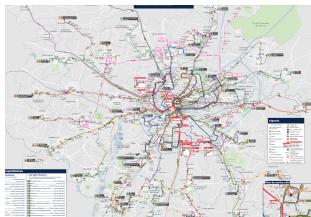
*Gauthier LYAN*

## Contexte et motivations

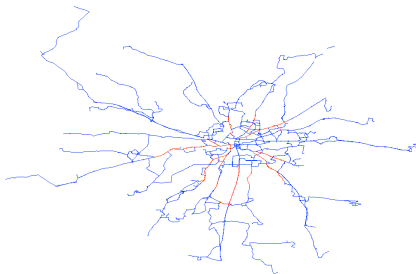
- ▶ Emergence des **smart-cities**
- ▶ **Masses de données** et open data
- ▶ Intérêt des **jumeaux numériques**
  - ▷ Analyse
  - ▷ Surveillance
  - ▷ Prédiction (what-if)
  - ▷ Couplage permanent au monde réel (ex. *Kubik building*, 3400 capteurs, bâtiment configurable, R&D efficacité énergétique)
- ▶ Virtualisation des villes et de leurs services



- ▶ Réseaux de transports **complexes** et **connectés**
- ▶ Comprendre, planifier, simuler le réseau en **multi-échelles**
- ▶ Comprendre les variations de la **vitesse commerciale** : vitesse des usagers
- ▶ **Prédire** le comportement de **nouvelles lignes de bus**
- ▶ **Meilleurs services pour la métropole**



Le réseau **STAR** de Rennes Métropole



# Verrous

---

- 🔒 **Peu d'accès** et **mauvaise qualité** des données [Robinson et al.(2014)]
- 🔒 **Coût de l'expertise métier** propre aux transports [Bouyekhf et al.(2003)]
- 🔒 Difficulté de prédiction **multi-échelles** [Mendes-Moreira et al.(2020)]
- 🔒 Chaîne logicielle **incomplète** ne passant pas à l'échelle

## Verrous

---

- 🔒 **Peu d'accès** et **mauvaise qualité** des données [Robinson et al.(2014)]
- 🔒 **Coût de l'expertise métier** propre aux transports [Bouyekhf et al.(2003)]
- 🔒 Difficulté de prédiction
- Multi-échelles** [Mendes-Moreira et al.(2020)]
- 🔒 Chaîne logicielle **Incomplète** ne passant pas à l'échelle

### Thèse

*“Dans le domaine des transports publics, réaliser des prédictions multi-échelles est possible par une intégration des données dans des modèles logiciels spatio-temporels couplés à de l'apprentissage automatique.”*

# Plan

---

1. Etat de l'art
2. Données : Préparation et usages chez Keolis Rennes
3. Compred : PRÉDictions COMPositionnelles
4. DataTime : Un métamodèle pour intégrer données, temps et espace
5. Conclusion

## Etat de l'art - Transports publics

---

- ▶ **Problèmes de qualité et d'accès aux données** [Smith et al.(2003)]
  - ▷ Multiplicité et variété des capteurs
  - ▷ Masses de données hétérogènes
  - ▷ Prégance de données erronées
- ▶ **Réseaux de transports = problème systémique** [Farahani(2013)]
  - ▷ Réseaux intriqués dans les métropoles (trafic, piétons, voirie, ...)
  - ▷ Optimisation à plusieurs variables (vitesse, pollution, nombre de voyageurs, ...)
  - ▷ Expertise et mesures nécessaires pour l'optimisation des réseaux

## Trois types d'outils

---

1. Modèles **mathématiques**
2. Modèles d'**apprentissage automatique**
3. Modèles de **simulation**



## Critères d'évaluation

---

- ▶ Besoin en expertise métier
- ▶ Usage de sources de données réelles et/ou hétérogènes
- ▶ Modèle spatio-temporel
- ▶ Prédications multi-cibles
- ▶ Evolutivité
- ▶ Passage à l'échelle
- ▶ Transposable à d'autres problématiques

## Etat de l'art - Analytique des transports publics

---

► **Modèles mathématiques** [Fernandez and Valenzuela(2003)]

$$VC = (V_0 + V'_0 \delta D + V''_0 \delta P + V'''_0 \delta L) e^{-[(\alpha + \alpha' \delta fp)fp + (\beta + \beta' \delta tp)tp + (\gamma + \gamma' \delta fs)fs + (\sigma + \sigma' \delta ts)ts]} \quad (1)$$

- ✓ Passage à l'échelle
- ✗ Expertise métier nécessaire
- ✗ Aucun usage de données
- ✗ Fixe dans l'espace
- ✗ Dédié à un modèle précis (vitesse, temps, consommation, ...)

# Etat de l'art - Analytique des transports publics

## ► Modèles Machine Learning [Altinkaya and Zontul(2013)]

- ✓ Usage de sources hétérogènes
- ✓ Peu d'expertise métier
- ✗ Pas évolutif

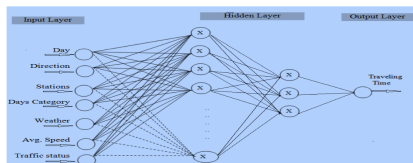


Figure 2: Proposed neural network structure

Réseau de neurones tel que conçu par [Zaki et al.(2013)]

# Etat de l'art - Analytique des transports publics

## ► Simulations [Cats et al.(2010)]

- ✓ Modèle spatio-temporel
- ✓ Prédictions multi-cibles
- ✗ Expertise métier requise
- ✗ Pas de couplage aux données réelles
- ✗ Dédié à une seule problématique
- ✗ Pas de passage à l'échelle et peu évolutif

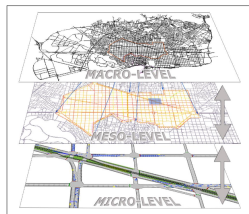
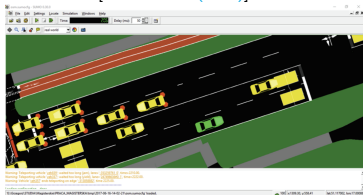


Figure 2: From Macro to Meso and Micro

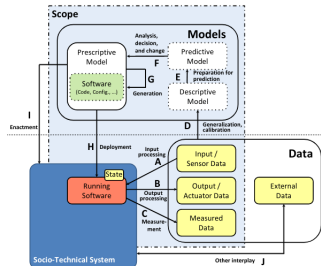
Illustration de simulation multi-échelle par [Barceló et al.(2005)]



Exemple d'usage de l'outil de simulation SUMO par [Ciskowski et al.(2019)]

# Etat de l'art - Génie logiciel

- ▶ Ingénierie dirigée par les modèles [Combemale et al.(2020)]
  - ✓ Pose les bases de concepts **MO**dels and **DA**ta
  - ✗ Pas spatio-temporel
- ▶ Analytique de graphes temporels à évolution rapide [Hartmann et al.(2019)]
  - ✗ Modèle à évolution rapide
  - ✗ Méthode de prédiction non spécifiée
  - ✗ Propriétaire



Le Metamodèle MODA par [Combemale et al.(2020)]

# Positionnement

Caractéristiques	Simu.	Maths.	ML	Cette thèse
Sans expertise métier	✗	✗	✗	✓
Sources hétérogènes	✗	✗	✓	✓
Données réelles	✗	✗	~	✓
Spatio-temporel	✓	✗	~	✓
Predictions multiples	✓	✗	~	✓
Evolutif	✗	~	✗	✓
Passé à l'échelle	✗	✓	~	✓
Transposable	✗	✗	~	✓

Obtenir un meta modèle temporel et centré sur un graphe commun permettant par apprentissage automatique des prédictions multi-échelles en couplage permanent avec les données réelles.

# Plan

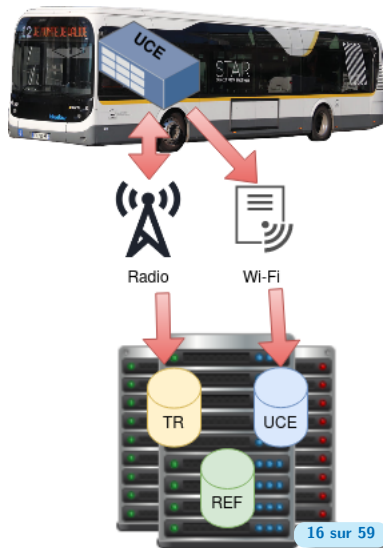
---

1. Etat de l'art
2. Données : Préparation et usages chez Keolis Rennes
3. Compred : PRÉDictions COMPositionnelles
4. DataTime : Un métamodèle pour intégrer données, temps et espace
5. Conclusion

## Données : Préparation et usages chez Keolis Rennes

### ► Existant

- Données embarquées (UCE)
- Données temps-réel (TR)
- Données de références (REF)





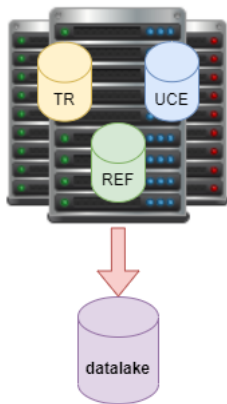
## ► Conception d'un datalake

- ▷ Supérieur à 100 Go
- ▷ ~ 15 batchs mensuels Apache Spark
- ▷ ~ 5000 lignes de code
- ▷ 9 mois de développement

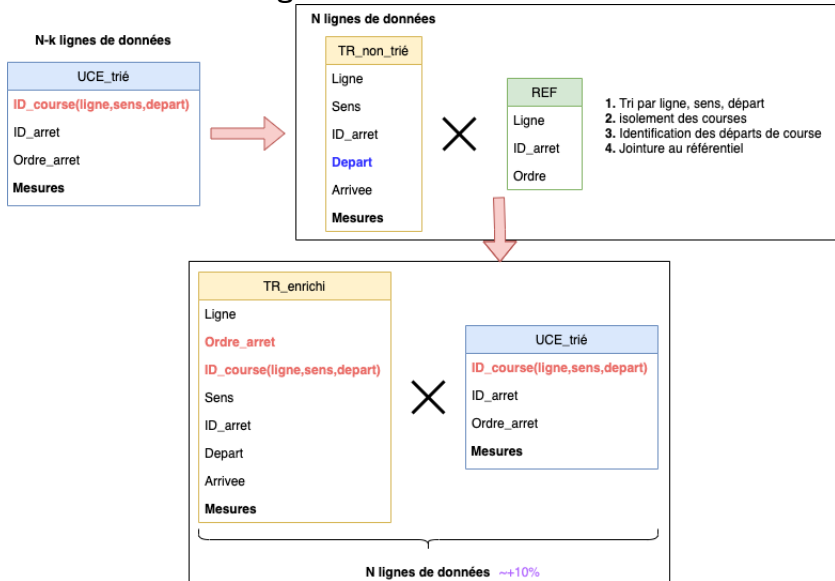
## ► Homogénéisation des sources de données

- ▷ Extraction des silos
- ▷ **Pivot d'intégration : Graphe(arrêts, inter-arrêts)**
- ▷ Qualification : règles de gestion (quelques minutes), consolidation (plusieurs heures)

## ► **Tâche** : Prédiction de la vitesse commerciale



# Méthodologie de consolidation des données



# Qualité de la prédiction de la VC selon le niveau de curation

---

▶ **4 jeux de données :**

1. **H0** : UCE brut
2. **H1** : UCE nettoyé (règles de gestion)
3. **H2** : (UCE U TR) brut ( $\sim +10\%$ )
4. **H3** : (UCE U TR) nettoyé (règles de gestion)

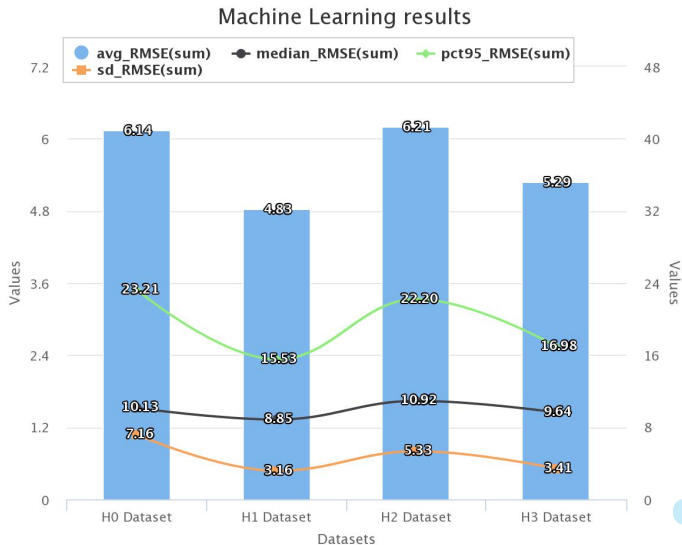
▶ **Mesure considérée RMSE<sup>1</sup>** de la prédiction de la VC.

---

1. Root Mean-Square Error : Racine carrée de la moyenne des erreurs carrées

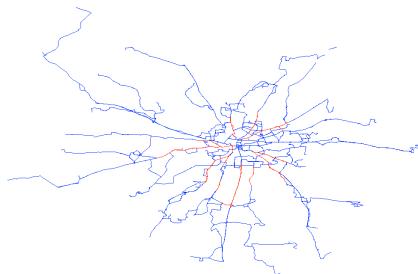
# Qualité de la prédiction de la VC selon le niveau de curation

- ▶ Données utilisables, résultats satisfaisants
- ▶ **Surqualifier les données est inutile**



- ▶ **Publication** : G. Lyan, D. Gross-Amblard, J-M. Jézéquel, S. Malinowski, «Impact Of Data Cleansing For Urban Bus Commercial Speed Prediction», *Springer Nature Computer Science. In press 2021.*

- ▶ **Analyse d'impact VC**
  - ▷ **Croisement de sources externes** (trafic, météo, vélos, fréquentation)
  - ▷ Extraction de **données OpenStreetMaps (OSM)**
    - ▶ Encadrement de **2 stagiaires** : Oklin KONE et Fabrice TRA
  - ▷ Aboutissement : **extracteur OSM open source, transposable sur d'autres métropoles**, 1000 lignes de code.



Représentation visuelle des données **OSM** et **open trafic** correspondantes avec le réseau de Rennes (STAR)

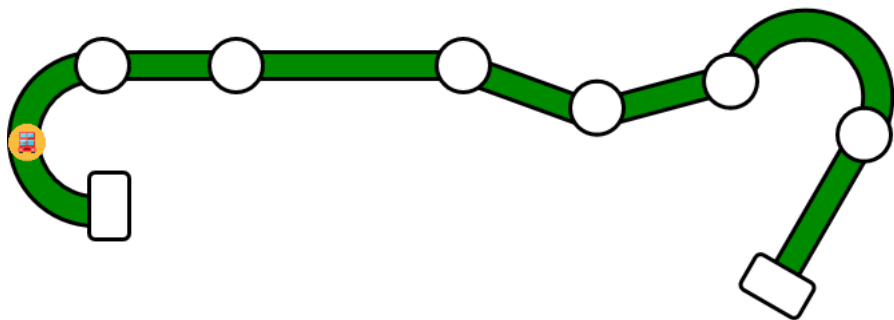
- ▶ Mise en évidence des **facteurs influants la VC à Rennes** :
  - ▷ Impact des feux (-0.33)
  - ▷ Impact du trafic routier (-0.25)
  - ▷ Impact de la fréquentation et du mode de validation [-0.2; -0.1]
  - ▷ Impact des feux dans les couloirs bus (- 0.13)
  - ▷ Impact des vélos (*mesures sur site*)
  - ▷ Impact du mode de propulsion des bus (*mesures terrain*)
- ▶ **Cohérence des résultats avec l'état de l'art sur d'autres réseaux** [Fernandez and Valenzuela(2003), Hofmann and O'Mahony(2005), Jäppinen et al.(2013), Cortés et al.(2011), Tirachini(2013)]

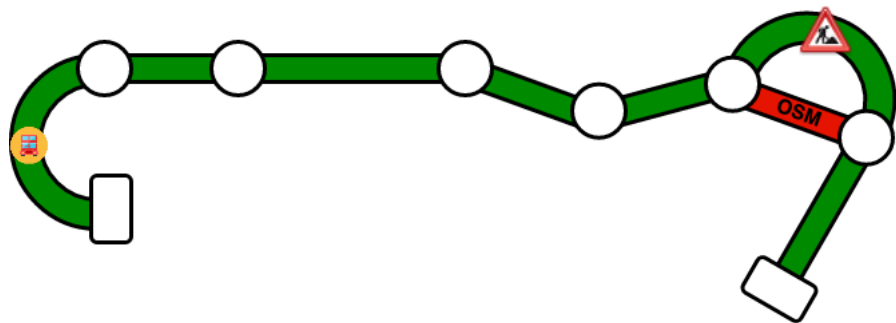
# Plan

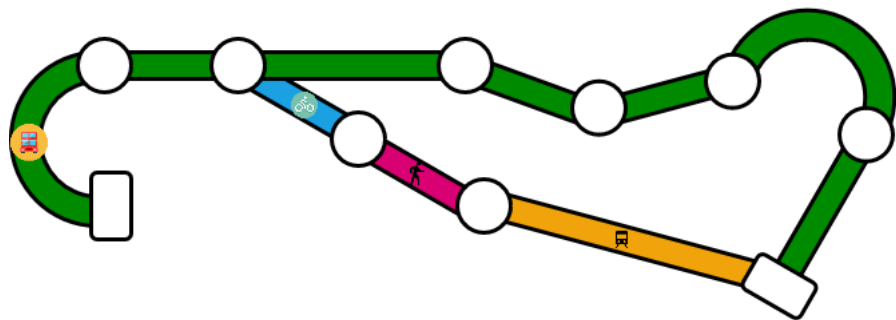
---

1. Etat de l'art
2. Données : Préparation et usages chez Keolis Rennes
3. **Compred : PRÉDictions COMPositionnelles**
4. DataTime : Un métamodèle pour intégrer données, temps et espace
5. Conclusion









# Compred : PRÉDictionS COMPositionnelles

- ✓ Données identifiées et qualifiées
- ✓ Apprentissage automatique possible
- ✓ Basé sur modèle graphe
- ✓ **Un seul modèle ML pour l'ensemble du réseau**
  - ▶ Modèle **bottom-up** (grain fin vers grain large)
  - ▶ Lois de **composition** (somme ou produit)
  - ▶ Prédiction **multi-échelles de qualité ?**

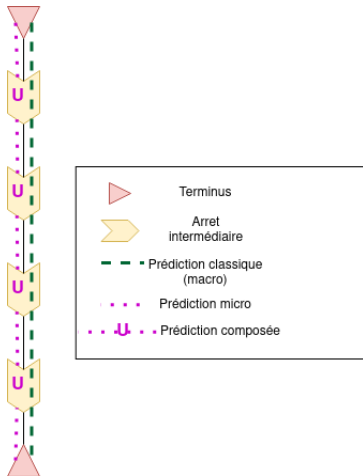


Illustration de la prédiction compositionnelle vs prédiction "classique"

## Méthodologie

- ▶ **Proposition de 2 lois de composition** :  $\Sigma$  (Temps de parcours) et  $\Pi$  (Accidentologie)
- ▶ **Jeu de données** : (UCE U **TR**) U **OSM** sur l'année **2019**
- ▶ **Sélection** de **13 lignes** de bus (chrono, metrop, urbaines)

## Pour chacune des lignes

- ▶ **Jeu d'apprentissage** = Différence du **dataset total** et des **données de la ligne** de bus
- ▶ **Jeu de test** = Jeu de données de la ligne de bus
- ▶ **Tâche** : Prédiction de la VC de la ligne de test
  - ▷ **Macro** : `learn(chemins) | predict(nouveaux chemins)`
  - ▷ **Notre approche** : `learn(arcs) | composition(pred(arc1), pred(arc2), ..., pred(arcN))`

## Validation

- ▶ **Application** de la méthode sur les **données réelles**
- ▶ Comparaison des résultats **Macro** vs **Composée**

## Résultats des prédictions de VC des lignes testées

Méthode	RMSE*	MAE*	MAPE*
<b>Macro</b>	3.2 km/h	2.9 km/h	12.2%
<b>Composition</b>	3.0 km/h	2.4 km/h	12.1%
<i>Yidan et al.</i>	3.4 km/h	2.1 km/h	14.8%

Résultats Macro et Composés sur données réelles en comparaison avec l'état de l'art récent [[Yidan et al.\(2019\)](#)]

- ✓ **Conclusion** : Prédictions de qualité et **multi-échelles**
- ✓ **Conclusion** : **Résultats composition au niveau de l'état de l'art**
- ✓ **Valide également** sur des jeux de données artificiels (vitesse et viabilité)
- ✓ **Publication** : G. Lyan, D. Gross-Amblard, J-M. Jézéquel, «On the Quality of Compositional Prediction for Prospective Analytics on Graphs», *MLKGraphs workshop 2021, Accepted 2021*.

\* *Indicateurs de précision classiques des méthodes d'apprentissage automatique*



# Plan

---

1. Etat de l'art
2. Données : Préparation et usages chez Keolis Rennes
3. Compred : PRÉDictions COMPositionnelles
4. DataTime : Un métamodèle pour intégrer données, temps et espace
5. Conclusion

## DataTime : Un métamodèle pour intégrer données, temps et espace

---

- ✓ Modèle spatio-temporel
- ✓ Modèle prédictif mutli-echelles

**To-Do** Intégrer les séries temporelles au modèle graphe

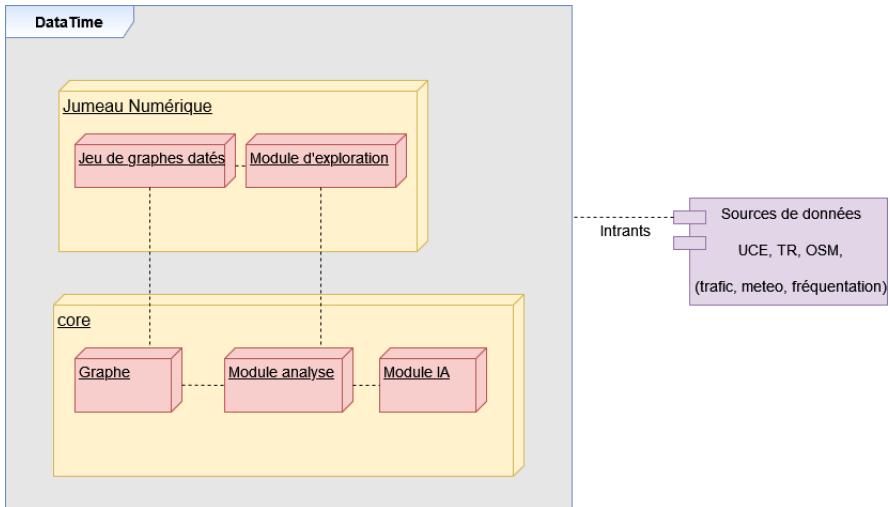
**To-Do** Implémenter les fonctions de machine learning

**To-Do** Implémenter les lois de composition

**To-Do** Implémenter l'exploration des données

## Proposition

- ▶ Intégration des données au sein d'un modèle logiciel graphe multi-échelle
- ▶ Intégration logicielle de modèles ML
- ✓ DataTime : **Analyse intégrée de données passées, présentes ou futures (prédites)**



Meta-architecture de DataTime

## Exemples de code

---

### Sélection d'une ligne de bus

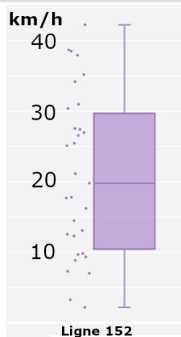
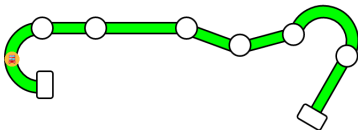
```
1: val busLine :Line = KeolisGraph.getLine("152", Direction.B)
```

### Analyse de données historiques

```
1: /** TimePeriod pour laquelle des données historiques existent **/  
2:  
3: var tp :TimePeriod = TimePeriod("2019-01-01 08 :00 :00","2021-02-01  
   00 :00 :00")  
4:  
5: var metric :Metric = Speed/** TravelTime, Freq**/  
6:
```

## Analyse de données historiques

- 1: `/** Renvoie toutes les courses trouvées dans les données **/`
- 2:
- 3: `var runs :Set[Run] = getMetric(busLine,tp,metric)`
- 4: `runs.displayMetric()`

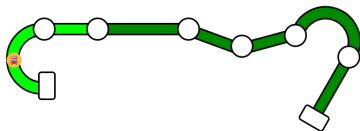


## Analyse de données temps-réel

```
1: /** TimePeriod pour des données temps réel */  
2:  
3: val now :Long = System.currentTimeMillis  
4: tp.setPeriodStart(now)  
5: tp.setPeriodEnd(now)
```

## Analyse de données temps-réel

- 1: `/** Renvoie les données de la première section de la dernière course en date */`
- 2:
- 3: `runs = getMetric(busLine.getSections.head,tp,metric)`
- 4: `runs.displayMetric()`



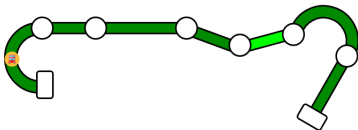


## Prédiction pour une ligne existante

```
1: /** TimePeriod pour une période future = pas de données historiques **/  
2:  
3: tp.setPeriodStart("2022-03-19 17 :30 :00")  
4: tp.setPeriodEnd("2022-03-19 17 :30 :00")
```

## Prédiction pour une ligne existante

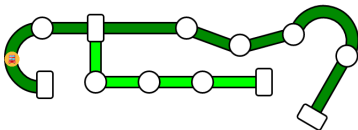
- 1: `/** Renvoie un temps de parcours prédit sur le 5ème inter-arrêt de la ligne */`
- 2:
- 3: `runs = getMetric(busLine.getInterStation(5),tp,metric)`
- 4: `runs.displayMetric()`





## Prédiction pour une nouvelle ligne

- 1: /\*\* Renvoie un résultat de fréquentation prédit pour la nouvelle ligne de bus \*\*/
- 2:
- 3: runs = getMetric(**nline**,tp,metric)
- 4: runs.displayMetric()



## Résultats

---

- ✓ **Analyse spatio-temporelle** du réseau STAR
- ✓ Capacité de prédiction à **n'importe quelle échelle spatio-temporelle**
- ✓ **Aucune exigence** en connaissance métier
- ✓ Utilisation de **sources de données hétérogènes**
- ✓ **Multiples cibles** de prédiction
- ✓ **Transposable** à d'autres problématiques (réseaux de transport, smart-grids, ...)
- ✓ **Evolutif** (ajout de sources, de modèles, création de chemins, etc.)
- ✓ **Passage à l'échelle** : Capacité à monter en charge sur les architectures distribuées
- ✓ **Libère l'exigence métier pour d'autres tâches chez Keolis Rennes**

- ✓ **Publication** : G. Lyan, J-M. Jézéquel, D. Gross-Amblard, B. Combemale, «DataTime : a Framework to smoothly Integrate Past, Present and Future into Models», *MODELS 2021*, *Accepted 2021*.

# Plan

---

1. Etat de l'art
2. Données : Préparation et usages chez Keolis Rennes
3. Compred : PRÉDictions COMPositionnelles
4. DataTime : Un métamodèle pour intégrer données, temps et espace
5. Conclusion

## Conclusion et perspectives

Caractéristiques	Simu.	Maths.	ML	Cette thèse
Sans expertise métier	✗	✗	✗	✓
Sources hétérogènes	✗	✗	✓	✓
Données réelles	✗	✗	~	✓
Spatio-temporel	✓	✗	~	✓
Predictions multiples	✓	✗	~	✓
Evolutif	✗	~	✗	✓
Passé à l'échelle	✗	✓	~	✓
Transposable	✗	✗	~	✓

- ▶ **Validation et généralisation** des facteurs influents la vitesse des bus via les masses de données
- ▶ **DSL** pour DataTime et extension des sources de données
- ▶ **Meta-learning intégré à DataTime** pour optimiser le choix de modèle ML



# Merci de votre attention

---

## Rappel des contributions

- ▶ Un **datalake** pour **Keolis**
  - ✓ **Publication sous presse chez Springer Nature Computer Science**
- ▶ **Identification** des **impacts** sur la **vitesse commerciale** des bus à Rennes
- ▶ **Prédictions** compositionnelles : **Multi-échelles**
  - ✓ **Publication sous presse chez MLKGraphs 2021**
- ▶ DataTime : **Un cadre logiciel pour intégrer données, passé, présent et futur**
  - ✓ **Publication sous presse chez MODELS 2021**

# Bibliographie (1)

---



Mehmet Altinkaya and Metin Zontul.

Urban Bus Arrival Time Prediction : A Review of Computational Models.  
2(4) :7, 2013.



et al. Barceló et al.

A methodological approach combining macro, meso and micro simulation models for transportation analysis.  
page 24, 2005.



R. Bouyekhf et al.

Modelling, performance evaluation and planning of public transport systems using generalized stochastic Petri nets.

*Transport Reviews*, 23(1) :51–69, January 2003.

ISSN 0144-1647, 1464-5327.

doi : 10.1080/01441640309898.

URL <http://www.tandfonline.com/doi/abs/10.1080/01441640309898>.

## Bibliographie (2)

---



Oded Cats et al.

Mesoscopic Modeling of Bus Public Transportation.

*Transportation Research Record : Journal of the Transportation Research Board*, 2188(1) :9–18, January 2010.

ISSN 0361-1981, 2169-4052.

doi : 10.3141/2188-02.

URL <http://journals.sagepub.com/doi/10.3141/2188-02>.



Piotr Ciskowski et al.

Estimation of Travel Time in the City Using Neural Networks Trained with Simulated Urban Traffic Data.

In Wojciech Zamojski, Jacek Mazurkiewicz, Jarosław Sugier, Tomasz Walkowiak, and Janusz Kacprzyk, editors, *Contemporary Complex Systems and Their Dependability*, volume 761, pages 121–134. Springer International Publishing, Cham, 2019.

ISBN 978-3-319-91445-9 978-3-319-91446-6.

doi : 10.1007/978-3-319-91446-6\_13.



Benoit Combemale et al.

A Hitchhiker's Guide to Model-Driven Engineering for Data-Centric Systems.

*IEEE Software*, 2020.

ISSN 0740-7459, 1937-4194.

doi : 10.1109/MS.2020.2995125.

## Bibliographie (3)

---



Cristian E. Cortés et al.

Commercial bus speed diagnosis based on GPS-monitored data.

*Transportation Research Part C : Emerging Technologies*, 19(4) :695–707, 8 2011.

ISSN 0968090X.

doi : 10.1016/j.trc.2010.12.008.



Reza Zanjirani Farahani.

A review of urban transportation network design problems.

*European Journal of Operational Research*, page 22, 2013.



R. Fernandez and E. Valenzuela.

A model to predict bus commercial speed.

*Traffic Engineering & Control*, 44(2), 2 2003.

ISSN 0041-0683.



Thomas Hartmann et al.

GreyCat : Efficient what-if analytics for data in motion at scale.

*Information Systems*, 83 :101–117, 7 2019.

ISSN 03064379.

doi : 10.1016/j.is.2019.03.004.

## Bibliographie (4)

---



Markus Hofmann and Margaret O'Mahony.

The Impact of Adverse Weather Conditions on Urban Bus Performance Measures – an Analysis Using ITS Technology.

In *Intelligent Transportation Systems*, page 7. IEEE, 2005.



Sakari Jäppinen et al.

Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki : An open data approach.

*Applied Geography*, 43 :13–24, 9 2013.

ISSN 01436228.

doi : 10.1016/j.apgeog.2013.05.010.



João Mendes-Moreira et al.

Reconciling Predictions in the Regression Setting : An Application to Bus Travel Time Prediction.

In Michael R. Berthold, Ad Feelders, and Georg Kreml, editors, *Advances in Intelligent Data Analysis XVIII*, volume 12080, pages 313–325. Springer International Publishing, Cham, 2020.

ISBN 978-3-030-44583-6 978-3-030-44584-3.

doi : 10.1007/978-3-030-44584-3\_25.

URL [http://link.springer.com/10.1007/978-3-030-44584-3\\_25](http://link.springer.com/10.1007/978-3-030-44584-3_25).

Series Title : Lecture Notes in Computer Science.

## Bibliographie (5)

---



Steve Robinson et al.

Methods for pre-processing smartcard data to improve data quality.

*Transportation Research Part C : Emerging Technologies*, 49 :43–58, December 2014.

ISSN 0968090X.

doi : 10.1016/j.trc.2014.10.006.

URL <https://linkinghub.elsevier.com/retrieve/pii/S0968090X14002988>.



Brian Smith et al.

Exploring Imputation Techniques for Missing Data in Transportation Management Systems.

*Transportation Research Record : Journal of the Transportation Research Board*, 1836 :132–142, 1 2003.

ISSN 0361-1981.

doi : 10.3141/1836-17.



Alejandro Tirachini.

Bus dwell time : the effect of different fare collection systems, bus floor level and age of passengers.

*Transportmetrica A : Transport Science*, 9(1) :28–49, 1 2013.

ISSN 2324-9935, 2324-9943.

doi : 10.1080/18128602.2010.520277.

# Bibliographie (6)

---



Sun Yidan et al.

*Bus Travel Speed Prediction using Attention Network of Heterogeneous Correlation Features.*

Society for Industrial and Applied Mathematics, Philadelphia, PA, May 2019.

ISBN 978-1-61197-567-3.

doi : 10.1137/1.9781611975673.



M Zaki et al.

Online Bus Arrival Time Prediction Using Hybrid Neural Network and Kalman filter Techniques.

3(4) :7, 2013.

## Détail des mesures

### RMSE

Racine carrée de l'erreur quadratique moyenne

$$RMSE = \sqrt{\frac{1}{N} \sum_1^N (Y_t - \hat{Y}_t)^2} \quad (2)$$

### MAE

Erreur moyenne absolue

$$MAE = \frac{\sum_1^N (Y_t - \hat{Y}_t)}{N} \quad (3)$$

### MAPE

Pourcentage absolu d'erreur moyenne

$$MAPE = \frac{1}{N} \sum_1^N \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| * 100$$



## Détail de l'équation de [Fernandez and Valenzuela(2003)]

$$VC = (V_0 + V_0' \delta D + V_0'' \delta P + V_0''' \delta L) e^{-[(\alpha + \alpha' \delta f_p) f_p + (\beta + \beta' \delta t_p) t_p + (\gamma + \gamma' \delta f_s) f_s + (\sigma + \sigma' \delta t_s) t_s]} \quad (5)$$

Variable	Description
$V_c$	Vitesse commerciale moyenne (km/h)
$f_p$	Nombre d'arrêts de bus par km
$f_s$	Nombre d'arrêts aux feux de circulation
$t_p$	Temps de passage aux arrêts en secondes
$t_s$	Temps d'attente aux feux de circulation en secondes
$\delta D$	Variable de direction $\in [0, 1]$
$\delta P$	Variable d'heure de pointe $\in [0, 1]$
$\delta L$	Variable de voie bus $\in [0, 1]$
$\delta f_x, \delta t_x$	Variables de boîte auto $\in [0, 1]$
$V_0^i, \alpha^i, \beta^i, \gamma^i, \sigma^i$	Variables de calibration

## Résultats compred synthétiques

---

Modèle	RMSE	MAE	MAPE
<b>Temps de parcours (somme de temps)</b>			
<b>Micro</b>	4.78s	0.12s	7.49%
<b>Macro</b> (Bus de ligne entière)	62.36s	51.96s	3.76%
<b>Compred</b> (Composition des prédictions)	25.36s	3.23s	1.75%
<b>Viabilité (produit de probabilités)</b>			
<b>Micro</b>	0.019	0.001	1.50%
<b>Macro</b> (Bus de ligne entière)	0.073	0.012	7.69%
<b>Compred</b> (Composition des prédictions)	0.072	0.006	7.63%

# Données synthétiques

$$\text{TravelTime} = 35 * e^{-(0.03*d+0.02*h+0.05*p+0.0002*l+0.002*s+0.001*c+0.0002*g+0.0003*ra+0.007*st+0.01*ts+0.0002*pb-0.0001*pc+0.0001*po+0.0002*ps)} \quad (6)$$

$$\text{Risk} = 1 * e^{-(0.00002*d+0.00002*h+0.000035*p+0.000002*l+0.000005*s+0.000003*c+0.000005*g+0.00003*ra+0.000002*st+0.000005*ts+0.000002*pb-0.000001*pc+0.000001*po+0.000002*ps)} \quad (7)$$

Variable	Meaning
d	jour de la semaine $\in [0, 6]$
h	vacances $\in [0, 1]$
p	tranche horaire $\in [0, 2]$ 0=libre, 1=trafic moyen 2= trafic fort
l	longueur en mètres $\in [150, 600]$
s	vitesse légale $\in [30, 55]$
c	nombre de passages piétons $\in [0, 4]$
g	nombre de cédez-le-passage $\in [0, 3]$
ra	nombre de rond-points $\in [0, 2]$
st	nombre de stops $\in [0, 1]$
ts	nombre de feux $\in [0, 3]$
pb	part de route partagée vélo $\in [0.0, 1.0]$
pc	part de voie bus $\in [0.0, 1.0]$
po	part de route en sens unique $\in [0.0, 1.0]$
ps	part de zone 30 $\in [0.0, 1.0]$

Données générées grâce à une loi normale :  $normal(\mu, \sigma)$ , avec  $\mu$  moyenne générée via les Equations 6 et 7, et  $\sigma$  l'écart type égal à  $\theta\mu$  avec  $\theta \in [0.025, 0.15]$